DOCUMENT RESUME

ED 406 225 SE 059 974

AUTHOR De Corte, Erik; And Others
TITLE Estimating Numerosity in a Rectangular Grid: The Development of a Clever Estimation Strategy.
PUB DATE Mar 97
NOTE 41p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Cognitive Development; Cognitive Psychology; Cognitive Style; College Students; *Computation; Constructivism (Learning); Discovery Processes; Elementary Education; *Estimation (Mathematics); Grade 2; Grade 6; Higher Education; Learning Strategies; *Numeracy; *Strategic Planning; Thinking Skills

ABSTRACT

This study investigates the development of an adaptive strategy for estimating numerosity from the theoretical perspective of strategic change. A simple estimation task was used. Participants of 3 different age groups (20 university students, 20 sixth-graders and 10 second-graders) had to estimate 100 numerosities of colored blocks presented in a 10 X 10 rectangular grid. This task allows for two distinct estimation procedures: (1) repeatedly adding estimations of groups of blocks (= addition procedure); or (2) subtracting the estimated number of empty squares from the (estimated) total number of squares in the grid (= subtraction procedure). A rational task analysis indicates that the most efficient overall estimation strategy consists of adaptive use of both procedures, depending on the ratio of blocks to empty squares. The first hypothesis was that there would be a developmental difference in the adaptive use of the two procedures. According to the second hypothesis, this adaptive use will result in better estimation accuracy. Converging evidence from different kinds of data (i.e., response times, error rates, and retrospective reports) supported both hypotheses. Contains 22 references. (Author/PVD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 406 225

ESTIMATING NUMEROSITY IN A RECTANGULAR GRID:

THE DEVELOPMENT OF A CLEVER ESTIMATION STRATEGY

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

E. De Corte

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Erik DE CORTE

Lieven VERSCHAFFEL

Christian LAMOTE

Nele DHERDT

Center for Instructional Psychology and Technology (CIP&T)

University of Leuven

Belgium

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

March 1997

Paper presented in a session on "Investigations of thinking, reasoning and strategy use in mathematics", organized at the Annual Meeting of the American Educational Research Association, Chicago, March 24-28, 1997

Address for correspondance: Prof. Dr. Erik De Corte, University of Leuven, Center for Instructional Psychology and Technology (CIP&T), Vesaliusstraat 2, B-3000 Leuven, Belgium - Tel. int-32-16-32.62.48 - Fax int-32-16-32.62.74 - E-mail: Erik.DeCorte@ped.kuleuven.ac.be

BEST COPY AVAILABLE

ABSTRACT

This study investigates the development of an adaptive strategy for the estimation of numerosity from the theoretical perspective of "strategic change" (Lemaire & Siegler, 1995; Siegler & Shipley, 1995). A simple estimation task was used in which participants of three different age groups (20 university students, 20 sixth-graders and 10 second-graders) had to estimate 100 numerosities of (colored) blocks presented in a 10 x 10 rectangular grid. Generally speaking, this task allows for two distinct estimation procedures: either repeatedly adding estimations of groups of blocks (= addition procedure) or subtracting the estimated number of empty squares from the (estimated) total number of squares in the grid (= subtraction procedure). A rational task analysis indicates that the most efficient overall estimation strategy consists of the adaptive use of both procedures, depending on the ratio of the blocks to the empty squares. The first hypothesis was that there will be a developmental difference in the adaptive use of the two procedures, and according to the second hypothesis this adaptive use will result in better estimation accuracy. Converging evidence from different kinds of data (i.e., response times, error rates, and retrospective reports) supported both hypotheses. From a methodological point of view, the study shows the potential of Beem's (1995a, 1995b) "segmentation analysis" for unravelling subject's adaptive choices between different procedures in cognitive tasks, and for examining the relationship between these adaptive choices and performance.

THEORETICAL BACKGROUND

Based on a review of the literature on strategic change, Lemaire and Siegler (1995) conclude that "usually, when the contribution of strategic change to learning is recognized at all, it is viewed solely in terms of acquisition of new, more effective strategies" (p. 94). In Siegler et al.'s own longitudinal studies of children's acquisition of single-digit addition and multiplication skills (Lemaire & Siegler, 1995; Siegler & Jenkins, 1989; Siegler & Shrager, 1984), it was found that learning these computational skills could not be explained in terms of this type of strategy change only. The substantial improvements in speed and accuracy observed among pupils derived also from three other sources of strategic change, relating to when each strategy was used, how each strategy was executed, and how (adaptively) strategies were chosen. Therefore, they developed a model that distinguishes four dimensions of strategic competence: changes in any one of those dimensions can yield overall improvements in speed and accuracy of computation: (a) acquisition of new strategies and the abandonment of old ones, (b) shift towards greater use of the more efficient available strategies, (c) improvement in the fluency and efficiency with which strategies are executed, and (d) increase of the adaptive nature of the choice among strategies. Siegler and Shipley's (1995) simulation model of strategy development - the so-called "adaptive strategy choice model" (ASCM) - yields a number of specific predictions about these four dimensions of strategic change, as well as about relations between earlier and later performance of individual children. While the simulation's predictions have already been tested in several studies involving different kinds of computational and other skills (for a review see e.g., Lemaire & Siegler, 1995) these investigations were limited in the sense that they concerned always learning to make adaptive choices between fact-retrieval and backup strategies, mostly at points when children were already quite proficient on the types of problems they were presented (Lemaire & Siegler, 1995, p. 95).

A first aim of the present study was to contribute to the further understanding of strategy choices in relation to learning, by focusing on the acquisition of adaptive choices in a task environment that has not been investigated before from the perspective of strategic choice and that was, moreover, unfamiliar to the participants, namely estimating numerosities presented in a rectangular grid. Besides contributing to an empirically-based theory of strategy change, this study had also a methodological aim, namely to test the potential of Beem's (1995) "program for

fitting two-phased segmented curve models with an unknown change point" for unravelling adaptive choices between different procedures in cognitive tasks and for examining the relationship between such adaptive choices and performance.

RATIONAL TASK ANALYSIS AND HYPOTHESES

Subjects in the present study were administered an estimation task consisting of 100 trials; each trial asked for the estimation of a numerosity of blocks presented in a rectangular grid with a length and a width of 10 square units. Each of the 100 little squares of this 10 x 10 grid were either empty or filled with a colored block; thus, the total number of blocks in the grid was always a number between 1 and 100. The task was always to estimate the total number of blocks (see Figure 1a and 1b for two example trials).

Insert Figure 1 about here

The particular geometrical structure in which the numerosities were presented, allows for the use of at least two distinct estipmation procedures: (1) an addition procedure, in which the subject divides the given quantity into a number of groups and then repeatedly adds the estimations of all these groups, or (2) a subtraction procedure, in which the estimation of the number of blocks is based on subtracting the (estimated) number of empty squares in the rectangular grid from the (estimated) total number of squares in that grid.

Based on this distinction between these two procedures for estimating a given numerosity, one can identify two general strategies for the estimation task as a whole: a stereotyped addition (or ADD) strategy and an adaptive minimization (or MIN) strategy. The sterotyped ADD strategy involves always applying the addition procedure, irrespective of the given number of blocks. For example, users of the ADD strategy will apply the same addition procedure to solve both estimation tasks given in Figure 1. In the MIN strategy one flexibly or adaptively applies the addition procedure for certain trials and the subtraction procedure for

other trials, the choice for an additive or a subtractive procedure being based on the ratio of blocks to empty squares. For instance, trials with few blocks and many empty squares -- such as the first trial in Figure 1 -- are solved by an addition procedure (e.g. "6 in the upper left corner, plus 4 in the bottom left corner, plus a group of about 15 blocks in the right part of the grid... this makes a total of about 25"); but for trials with many blocks and few empty squares -- such as the second trial in Figure 1 -- the estimation is made by counting or estimating the number of empty squares in the grid (e.g., "3 in the left upper corner, 4 in the left lower corner, then 5 in the upper right corner, and 3 others in the lower right corner, this makes a total of 15"), and subtracting this estimated number from the total of 100 (e.g., "100 minus 15 = 85").

A rational task analysis indicates that the adaptive MIN strategy is considerably more efficient than the stereotyped ADD strategy. Indeed, users of the MIN strategy will always select and apply the procedure that has the greatest advantage over the other in speed, accuracy, or both, and, therefore, they will make quicker and/or better estimations than subjects who always use the addition procedure, especially on tasks wherein the subtraction procedure will obviously result in shorter response times and more accurate estimations. A major problem with the MIN strategy, however, relates to the point at which the addition procedure is replaced by the subtraction procedure. From a purely mathematical point of view, one could argue that the locus of this change point is the smallest trial wherein the number of blocks exceeds the number of empty squares (i.e., the trial with 51 colored blocks presented in a 10 x 10 grid), but the rational task analysis suggests that the change point will not be totally determined by this objective task characteristic. This location will rather be based on the subject's own evaluation of when the more efficient but also more complex subtraction procedure becomes advantageous over the addition procedure in terms of speed, accuracy or both. And this "subjective" change point does not necessarily correspond to the first trial with the number of blocks exceeding the number of empty squares.

However, the rational analysis indicates also that the use and efficiency of the MIN strategy will depend on subjects' mastery of several knowledge elements and abilities (which are not required in the ADD strategy). First of all, users of the MIN strategy must be familiar with "rectangular area" as a situational model of multiplication; in other words, they must realize that the size of the grid can (easily) be calculated by multiplying the number of rows and the number of columns in the rectangular grid. Second, the product of this multiplication (i.e., 10 x 10 = .)

must be automatized, so that it can be retrieved from long-term memory instantaneously (rather than having to use a time-consuming backup procedure). Third, they must master a procedure for mentally calculating subtractions of the type 100 - x = ., which is required for subtracting the (estimated) number of empty squares from tot (estimated) total grid size. Finally, they must be able to identify for themselves the types of trials on which each procedure will have the greatest advantage over the other in speed, accuracy, or both, and also to plan and monitor their own estimation processes accordingly.

Starting from this a priori task analysis, the following two hypotheses were raised.

First, the use of the stereotyped ADD strategy will decrease and the use of the adaptive MIN strategy will increase with age. The following arguments underly this hypothesis. First, as said before, the MIN strategy requires mastery of several basic mathematical concepts (i.e., multiplication, subtraction) and skills (i.e., multiplication facts and the procedure for subtracting any number from 100) which may be absent or only partly acquired in young children (Fuson, 1992; Sowder, 1992; Verschaffel & De Corte, 1996). The second reason for hypothesizing greater use of the MIN strategy with increasing age, relates to the required metacognitive knowledge and monitoring skills (De Corte, Greer & Verschaffel, 1996). Indeed, the efficient use of the MIN strategy supposes the metacognitive understanding that there are constraints on the efficacy of always using the simple addition procedure, and that the information-processing demands of the task can be drastically decreased by making adaptive choices between the addition and the subtraction procedure. Research on metacognition suggests that the metacognitive knowledge needed for understanding the advantages of the MIN strategy and for applying it efficiently, is dependent of age (Schoenfeld, 1992). To test this first hypothesis, we compared the use of the ADD and the MIN strategy in three different age groups.

Based on the rational task analysis, we also hypothesized that -- regardless of age -- the use of the more flexible MIN strategy will result in higher performance on the estimation task. This second hypothesis was tested by analyzing the relationship between the use of the MIN strategy, on the one hand, and overall performance on the estimation task, on the other, in the three age groups.

To identify subjects' estimation strategies we analyzed the pattern of reaction times and estimation errors on all 100 trials by means of Beem's (1995a, 1995b) "program for fitting two-phased segmented curves models with an unknown change point", a computer program that

allows a more systematic and more fine-grained analysis of subjects' adaptive choices between different procedures than the more global and more rudimentary techniques for data-analysis used in previous research (see e.g., Newman & Berger, 1984). But, in addition we collected and analyzed other kinds of data, such as verbal reports and detailed analyses of subjects' systematic estimation errors and of their "direct hits" (= i.e. estimations that are exactly the same as the actual number of blocks in the grid rather than merely good estimations).

METHOD

Subjects

Fifty subjects participated in the experiment: 20 university students (majoring in psychology or educational sciences, and aged between 20 and 22 years), 20 sixth graders (aged between 11 and 12 years), and 10 second graders (aged between 7 and 8 years). The lower number of second graders was due to the fact that for practical reasons it was impossible to test these children during the school hours.

In the current mathematics curriculum in Flanders (Belgium) no explicit or systematic attention is spent at counting and estimating large numerosities (Verschaffel & De Corte, 1996). Therefore, it can be assumed that neither the subtraction procedure nor the MIN strategy has been systematically taught to the participants in this study. Concerning the conceptual and procedural underpinnings of the MIN strategy, the meaning of the basic arithmetic operations of addition, subtraction and multiplication is already introduced in the first grade of the primary school. In that grade pupils also learn how to add and subtract numbers up to 20. In the second grade the focus is on learning mental addition and subtraction in the number domain 20-100, and on learning the multiplication and division tables. At the time they were tested, the second graders had already received a first global introduction into the numbers up to 100 (e.g., they had learned to count up to 100); moreover they had some experience with simple mental additions and subtractions in this number domain (e.g., they had been solving sums like 24 + 10 = . , 36 - 8 = . by means of the so-called "hundred square", see Fuson, 1992), and they had learned the multiplication tables of 2 and of 3.

Materials

The estimation task was presented to the participants using an IBM-compatible 486 computer equipped with a VGA graphics card, with the resolution of the screen set to 640 x 480. Stimuli were rectangular grids consisting of 10 x 10 little square units. These square units could be either "on" or "off"; either the square was filled with a green-colored block or it remained empty (= in the same black color as the background of the whole screen display). Both the width and height of the grid were 200 pixels (about 9 cm on the screen). The square units were 20 by 20 pixels. The outline of the 10 x 10 rectangular grid was visible, and colored red. The generation of the 100 stimuli was completely randomized by the computer. The place of the blocks in the grid was also chosen at random. After each trial, the computer recorded the participants' answer as well as their response time.

Procedure

All participants were tested individually. They were told that 100 screen displays would be presented, each showing a number of green blocks, and that their task was to make an accurate estimation of the numerosities of blocks. An accurate estimation was defined as "something between a wild guess and the result of a precise counting and/or computation process" (Sowder, 1992; Treffers & De Moor, 1990; Verschaffel & De Corte, 1996). To help them get accustomed to the task and to understand what was meant by an accurate estimation, 10 example trials were given before the actual start of the experiment. During these 10 example trials, participants received immediate feedback about the accurateness of their estimation: a message appeared on the screen telling them whether their response was considered as an accurate or an inaccurate estimation. A 10 % deviation criterium was used to decide whether an estimation was considered accurate or not. This feedback was dropped as soon as the actual experiment started.

Each stimulus was presented for a maximum duration of 20 s. As soon as the participants had made their estimation, they had to press the space bar, which stopped the clock and blanked the screen. A new screen appeared which asked the participants to enter their estimation by means of the numerical keypad of the keyboard. When no answer was given

within 20 s., the screen blanked automatically, and the participant was asked to make an estimation.

At the end of the experiment, participants were briefly interviewed about the way they had handled the task. The major goal of this interview was to invite and help them to articulate the overall strategy they had used during the experiment. Each interview started with a general question about the participant's overall approach of the task. Participants who did not spontaneously and unambiguously report the use of the MIN strategy, were shown cards with screen prints of three different trials: a card with a small number of blocks (i.e. 27), a card with a very large number of blocks (i.e. 87), and a card with about half of the grid being filled with blocks (i.e. 54); for each card they were asked to describe as detailed as possible how they would solve that particular trial. During the interview, the experimenter did not make any suggestion about the (usefulness of the) MIN strategy. Answers were written down more or less verbatim.

There were some minor differences in the procedure for the three groups of participants. The second graders and the sixth graders had three small breaks during the test (after 25, 50, and 75 trials), while the adults made the 100 trials without any break.

Contrary to the sixth graders and the adults, the second graders were not asked to type their answers themselves; this was done by the experimenter. Try-outs had revealed that the procedure for typing in the answer was too distractive for these young children, and led to unreliable response-time data.

Finally, the introduction to the test was more extensive for the second graders than for the other two groups. Because of these young children's unfamiliarity with estimation of numerosity, the experimenter had to devote more time and effort to explain what a (good) estimation is and how it differs from blind guessing and precise counting.

RESULTS

Before presenting the results concerning the two research hypotheses, we present some general statistics concerning the response times and error rates for the three different age groups. Mean response times for the university students, the sixth graders and the second graders were 7111, 5902, and 7802 ms. Mean estimation error rates (= the mean overestimation or underestimation

of the number of blocks actually displayed on the screen) for these groups were 2.625, 4.717, and 8.621, respectively. Analysis of variance revealed that the factor "age group" did not have a significant effect on the response times, but had a significant effect on error rates [$F(2,46) = 38.5524$; $Ms_e = 3.0548$; $p<.0001$]. An a posteriori Tukey test revealed that all mean error rates differed from each other ($p<0.05$). In sum, the results showed no developmental differences in speed of estimation; but for estimation accuracy a clear developmental trend was found in the sense that older subjects made significantly better estimations than younger ones.

The discussion of the results concerning the hypotheses consists of three parts. First, we describe the development towards the MIN strategy (= hypothesis 1) as revealed by comparing the subjects' individual patterns of response times and estimation error rates for the 100 trials with the predictions derived from the two hypothetical strategies (ADD and MIN). To see which of these two strategies provides the best description of the individual response time and error rate data, linear regression lines as well as two-phase regression lines were fitted to the data. The second part contains the results related to the first hypothesis that are based on subjects' retrospective answers to the questions asked during the interview at the completion of the task, and on the analysis of systematic errors and "direct hits". In the third part the findings related to the second hypothesis are presented and discussed.

Fit of response times and error rates with a linear and a segmented curve model

The first and most important analysis of strategic development (= hypothesis 1) involved fitting the individual patterns of response times and the error rates with the predictions of the two hypothetical strategies (ADD and MIN).

This analysis required the removal of so-called "outliers", i.e. estimation errors which were "exceptionally" far removed from the correct numerosity due to a traceable erroneous estimation process. (Because these "outliers" yielded valuable additional evidence for the MIN strategy, they will be discussed in detail in the second part of the results section). Besides, we removed all cases wherein there was strong evidence for typing errors (e.g., responding with 887 while the correct answer was 87) or for incidentally premature pressing of the answer key (i.e. after less than 50 ms of presentation of the stimulus). This elimination procedure reduced

the datasets of the 20 adults and of the 20 sixth graders from 2000 to 1855 and 1970, respectively, and of the 10 second graders from 1000 to 992.

Beem's program for analyzing procedure shifts in cognitive tasks. As mentioned before, two types of statistical models were fitted to the individual response time and error rate data: a linear model and a two-phase segmented curve model.

Linear models are well known, and can be fitted using almost any statistics package. The relationship between the independent and the dependent variables -- in our case, respectively number of blocks and response time or error rate -- is described by a linear regression equation of the form y=a+bx. The parameters a and b, which are to be estimated, denote the intercept and the slope of the regression line.

The two-phase segmented curve model is less known and used, although it is -- according to Beem (1995a, 1995b; Ippel & Beem, 1987) -- ideally suited for the study of procedure shifts like those involved in the present study. Recently, Beem (1995a, 1995b) described an algorithm for fitting data to this kind of model. As the name "two-phase segmented curve" indicates, the relationship between the independent and the dependent variable is not described

$$\hat{y} = a_1 + b_1 x \ (x \leq s)$$
$$\hat{y} = a_2 + b_2 x \ (x > s)$$

by one but by two regression lines, which hold for different ranges of the independent variable. This model can be formally described as

For values of the independent variable up to s the first regression equation is used, while for values higher than s the second equation -- with a different intercept and slope -- holds. The parameter s is called the "change point" or "break point". Beem's program (1995a, 1995b) estimates the parameters and the change point using a least-squares algorithm, and always finds the global minimum of the error sum of squares. It should be noted that there exist two variants of the model: the join and the jump model. The first model is continuous in s. This means that the value of y at s is the same for both regression equations. When the model is not continuous in s, it is called a jump model. We will report the fits of the data with the jump model, since the

rational task analysis does not suggest continuity at the change point. Indeed, we anticipated that response times and error rates will be somewhat higher for the subtraction procedure, due to the extra steps involved in this complex procedure as compared to the simple addition procedure. As an illustration, we compare the trials with 25 and 75 blocks. In both cases a user of the MIN strategy has to count or to estimate a numerosity of 25: in the first case this numerosity of 25 involves the actual number of blocks, while in the second case this numerosity corresponds to the number of empty squares. While both counting or estimation processes will take about the same amount of time, in the latter case at least one important additional solution step has to be made, namely subtracting this numerosity of 25 from 100. Theoretically, this additional step should produce a longer response time and a greater error rate.

For users of the simple ADD strategy, we anticipated a relatively good fit of their response time and error rate data with the linear model, as the systematic application of the addition procedure will result in more or less linearly increasing response times and error rates as the number of blocks increases. For users of the MIN strategy, we expected a significantly better fit with the two-phase segmented curve model than with the linear model. Indeed, for these subjects both response times and error rates, will increase more or less linearly as long as they choose the addition procedure, but from a certain point on in the line of numerosities they will make the procedure shift from addition to subtraction, which will manifest itself in a more or less linear decrease in the pattern of response times and error rates. Thus, the crucial question is whether the segmented curve model provides a significantly better account of subjects' response time and error rate data than the simple linear model. The measure we used for this comparison between the alternative models, is $\underline{r}^2$ (or the proportion of explained variance).

To guarantee a proper interpretation of the results of the fitting data, the reader should keep in mind that a high $\underline{r}^2$ for the segmented curve model does not <u>necessarily</u> mean that the MIN strategy was applied. The estimated values of the parameters, and especially the slope parameters $\underline{b}_1$ and $\underline{b}_2$, need to be considered too, as the MIN strategy should lead to positive $\underline{b}_1$'s and negative $\underline{b}_2$'s. To illustrate the importance of these slope parameters, we refer to the (hypothetical) case of an user of the ADD strategy, whose reaction times continuously hit the ceiling of 20 s from a certain number of blocks on (e.g. from a numerosity of 50 on). The segmented curve model will explain this subject's response times far better than the linear model, but the absence of a negative $\underline{b}_2$ is a strong indication that he did not systematically use

the MIN strategy. Besides looking at the b_1 and b_2 parameters, qualitative inspections of the graphs of the individual response time and error rate data may be relevant too.

Response times patterns in the three age groups. Figure 2 presents the mean response times on the 100 trials in the three age groups.

Insert Figure 2 about here

To illustrate the diversity in the response time patterns in each group, Figure 3, 4 and 5 present the patterns of two students, two sixth graders and two second graders: one of a subject with a high fit with the segmented curve model (Figure 3a, 4a and 5a), and one of a subject with a low fit with that model (Figure 3b, 4b and 5b).

Insert Figure 3, 4 and 5 about here

The most important statistical data concerning the fit of the response times with the linear and the segmented curve model are given in Tables 1-3.

Insert Table 1, 2 and 3 about here

An inspection of the mean fits with the linear and the segmented curve model given in Tables 1-3, leads to the following observations.

First of all, for all subjects involved in the study the segmented curve model fits the data far better than the linear model. Application of Beem's (1995b) procedure to test the existence of

a change point, leads to the rejection of the null hypothesis of no change point - which would imply that the linear model is appropriate - for every subject (at the a=0.05 level).

Second, the r^2's for the segmented curve model increase from 27.4 for the second graders to 41 for the sixth graders and to 56.42 for the university students, while the r^2's for the linear model decrease from 8.32 for the second graders to 2.62 for the sixth graders to 0.81 for the university students. These observations are confirmed by a visual inspection of Figure 2: the response time pattern predicted by the MIN strategy is largely found among the university students, and to a lesser extent among the sixth graders, but for the second graders the pattern is completely different.

While these general observations are in line with the first part of the first hypothesis -- namely, that the use of the MIN strategy will increase with age -- they indicate at the same time that the ADD strategy is not the dominant strategy in the youngest age group. The results for the three groups will now be discussed separately.

As shown in Table 1, the university students' mean fit with the segmented model is 56.42. Interestingly, only a few subjects produced very high fits (i.e., fits of more than .80, like subject 3 of Figure 3a), but very low fits (i.e. of less than .50, like subject 12 from Figure 3b) are rare too.

The b_1 slopes are positive, indicating that up to a certain change point the response times increase as the number of blocks increases, and all b_2 slopes are negative, suggesting that after this change point the response times decrease as the number of blocks increases.

The mean value of the individual change points (= the point where the first regression equation ends and the second one begins) is 58.89. The finding that this change point is somewhat more than 50, corresponds with the rational task analysis stating that the subtraction procedure is more complicated than the addition procedure. Consequently, it is not surprising that the subjects of the oldest age group did not shift from the addition to the subtraction procedure as soon as the number of empty squares exceeded the number of blocks (i.e. at 51), but made this procedure shift somewhat later. Most individual change points lie rather close to this mean value; change points smaller than 50 and/or greater than 80 are almost completely absent.

Finally, a visual inspection of the graphs with the individual response time data (like those in Figure 3) reveals that for most university students the absence of a very high fit with the segmented curve model is not due to mismatches in the low and the high range of estimation values. For these trials students' data correspond very well to the predictions of the MIN strategy. But for the trials in the middle range, the response times are much noisier (see, for example, subject 12 in Figure 3b).

The results of the sixth graders are similar to those of the university students, although the mean $\underline{r}^2$ for the segmented curve model (41) is smaller and the mean $\underline{r}^2$ for the linear is model somewhat larger (2.62) (see Table 2). Moreover, compared to the university students, there are considerably less sixth graders with fits higher than 80 (for an exception, see subject 12 given in Figure 4a) and considerably more subjects with fits lower than 50 (see subject 16 given in Figure 4b).

Furthermore, some of the $\underline{b}_2$ slopes are very low (see, for instance, subjects 2, 5, 15) and for one sixth grader (i.e., subject 16 given in Figure 4b) the $\underline{b}_2$ slope is even positive.

The mean of the change points is 63.45, which is higher than for the university students (58.89). This suggests that the sixth graders shifted less quickly from the addition to the subtraction procedure than the students.

Finally, a visual inspection of the response time data reveals that many sixth graders have a middle range with very noisy response times, a range that is also larger than for most university students (see, for instance, the response time data of subject 16 in Figure 4b).

In sum, while the response time data of the sixth graders are still much better in accordance with the predicted pattern for the MIN strategy than with that for the ADD strategy, they suggest at the same time that the MIN strategy was applied less frequently and less systematically than among the university students.

The response time patterns of the second graders differ largely from those of the adults and the sixth graders (see Table 3). Compared to the two other age groups, the mean $\underline{r}^2$ for the segmented curve model is very low (27.4), while the $\underline{r}^2$ for the linear model is rather high (8.32). Remarkably, there are no second graders with fits larger than 50, and for some of them the $\underline{r}^2$ for the segmented curve model is not much higher than that for the linear model.

Moreover, only four out of the ten participants have negative b_2 slopes; this means that only for these four pupils there is a decrease in response time after the change point. Interestingly, the change point for these four children is very high (i.e., 86, 72, 87 and 87). This suggests that these pupils did use the subtraction procedure, but only for trials with a very small number of empty squares. An example of a data set of such a subject is given in Figure 5a.

For the six other second graders the b_2 slopes are positive. Most of them have a (very) small change point (i.e., lower than 40). Furthermore, they have (very) positive slopes in the low number range (before the change point), and almost flat slopes for the trials with medium and large numerosities (located behind that change point). Moreover, if we inspect the graphs of these six children (for an example, see Figure 5b), we see that the response time after this change point is typically rather low. (For some of these children, there is a number of numerosities at the very end of the line that show the linearly decreasing pattern that is typical for the subtraction procedure, but this number is too small to overrule the flat slope and to express itself in the statistical analysis.) Our interpretation of the approach followed by these six children is as follows: for items with small numerosities, they applied the addition procedure, or at least an "exact counting" variant of this estimation procedure. Once it became impossible for them to count all blocks in the available 20 s, these second graders began to use a kind of primitive and quick estimation procedure the duration of which was independent of the actual amount of blocks. Only for the trials with an (almost) completely filled grid, some of them exchanged this primitive estimation procedure for the subtraction procedure.

Taken as a whole, the response time patterns of the second graders contain considerably less evidence for the frequent and systematic use of the MIN strategy than those of the two other age groups. Nevertheless the response times suggest that many of them applied the subtraction procedure at least for trials with an almost completely filled grid, and that a minority of them used this procedure also on trials with numerosities from $\pm$ 80 on. Furthermore, although the second graders' fits with the segmented curve model are generally low, they are still considerably higher than those with the linear model, indicating that the hypothetical ADD strategy does also not provide a good account of the way in which these young children handled the estimation task.

Patterns of estimation error rates in the three age groups. As argued before, the ADD and the MIN strategies also predict different patterns of estimation error rates (= deviations of the participants' actual estimations from the correct ones). Whereas it is expected that the error rates of the ADD strategy users will increase as the numerosity of the blocks increases, users of the MIN strategy should have the highest error rates for numerosities from the middle range. Therefore, we compared the fit of the participants' actual error rates with those associated with the two hypothetical strategies in the same way as we did for the response time data.

Figure 6 presents a graphical representation of the mean error rates for the 100 trials in the three different age groups. A visual inspection of this figure reveals that - contrary to what was found for the response time data - the general pattern of error rates is similar in all three age groups. In each group the overall pattern is largely in line with the one predicted by the MIN strategy, namely relatively small error rates for low and high numerosities, and relatively high error rates for the numerosities in the middle range. However, this figure shows also that the absolute size of the error rates decreases strongly with age: while for the second graders the line of mean error rates climbs up to more than 15 units for the quantities from the middle range, it peaks at 10 and 5 for the sixth graders and the second graders, respectively.

The continuity in the overall pattern of error rates in the three age groups is confirmed by the statistics. For each group, the segmented curve model provides much better fits than the linear model: for the university students, the sixth graders, and the second graders the fit for the segmented curve model was, respectively, 40.57, 44.55, and 40 versus 0.51, 0.45, and 7.09 for the linear model. The mean change points for the three groups were 51.7, 51.7, and 58.9, respectively. Also, all b_1's were positive, and all but one of the b_2 slopes were negative (the one exception was from a second grader).

Thus, the error rate data confirm the conclusion from the analysis of the response time patterns, namely that the hypothetical MIN strategy provides a much better account of how the estimation task was handled than the ADD strategy. The most surprising result is that the decrease in the second part of the error rate patterns (behind the change point) is so clear in all age groups and that it begins so early, i.e. for numerosities of 50 to 60 already. This can be interpreted as an indication that the subtraction procedure was used more often than revealed by the response time data. In the next part of the results section, we will report the remaining results concerning the development of the adaptive MIN strategy.

BEST COPY AVAILABLE

Data from individual interviews, systematic estimation errors and "direct hits"

Individual interviews. As explained earlier, all participants were questioned about their overall approach to the estimation task during a short interview held at the end of the task. Based on these interview data, each subject was put into one of the following three categories: (a) participants who spontaneously reported that they had applied the MIN strategy; (b) participants whose verbal reports revealed the use of the MIN strategy only after they were shown the three cards (with 27, 54, and 87 blocks, respectively); (c) participants whose verbal reports did not contain any trace whatsoever of the use of the MIN strategy.

The analysis of the interview data revealed that almost all participants reported the use of the subtraction procedure for large numerosities in at least some cases. Thirteen out of the 20 university students spontaneously reported that they had applied the MIN strategy, and 6 others did so in reaction to the presentation of the three cards. Half of the sixth graders spontaneously explained that they had used the MIN strategy, while the other half did so in reaction to the presentation of the cards. Even in the group of second graders, nine out of 10 children produced verbal protocol data referring to the use of the subtraction procedure for at least some extreme trials: three second graders reported the use of the MIN strategy already at the beginning of the interview and six others did so after they were shown the cards. Although we are well aware of the methodological problems associated with the use and the interpretation of this kind of retrospective data (Ericsson & Simon, 1984), they nevertheless suggest that already at the age of 7 or 8 most subjects did not approach the estimation task in a stereotyped manner, but demonstrated at least some awareness of the existence of different procedures, and some skill in using these alternative procedures adaptively. The observed increase with age in the number of subjects who spontaneously reported the use of the MIN strategy can be considered as further confirming evidence for the more systematic -- but probably also more conscious use -- of this adaptive strategy in the older age groups.

Systematic estimation errors. The rational task analysis suggests that the MIN strategy may result in the following kinds of systematic and traceable estimation errors, which will be absent among users of the ADD strategy:

(1) errors based on an incorrectly determined total grid size, leading to systematic and large underestimations or overestimations for trials with large numbers of blocks eliciting the use of the subtraction procedure (= wrong total errors);

(2) errors due to the fact that -- at the end of the estimation process -- the subject has forgotten whether (s)he was engaged in an addition or a subtraction procedure, leading to estimation errors approximating to the number of empty squares instead of the actual numerosity of the blocks (= inversed errors).

A detailed qualitative analysis of the estimation errors revealed the occurrence of both types of errors, especially in the group of university students.

The first systematic error was observed in one university student. Like for most other subjects of his age group, this student's error rates increased gradually for numerosities up to about 70; once passed his change point of 70, the student suddenly and systematically began to make overestimations of about 70 and continued to do this for all remaining trials up to 100. The most plausible explanation for this strange pattern of estimation errors is that this student -- for some unclear reason -- had made an incorrect estimation of the total grid size at the beginning of the experiment (probably 13 x 13 = 169 instead of 10 x 10 = 100), and had continued to work with this wrong number each time he applied the subtraction procedure (i.e. for all trials with a numerosity between 70 and 100). Except for this single student, we found no other traces of the wrong total error among the subjects in the present study.

The second kind of systematic errors -- called "inversed errors" -- occurred more frequently. To be classified as an inversed error, the following two conditions had to be fulfilled: (a) the estimation deviates at least two standard deviations from the subject's mean error rate for all 100 trials, and (b) the estimation fits the so-called complement rule, stating that the sum of the subject's estimation and the actual number of blocks in the trial must equal 100 $\pm$ 5. Using these criteria, the number of inversed errors made by the students, the sixth graders and the second graders was 36 (= 1.9 %), 28 (= 1.4 %), and 6 (= 0.6 %), respectively. An ANOVA showed that this effect was significant [$F(2,46) = 11.01$; $MS_e = 1.67$; $p < .0462$]. A Tukey test ($a = 0.05$) revealed that the university students made significantly more inversed errors than the second graders.

In sum, the analysis of the systematic errors yields further evidence for the use of the MIN procedure in all groups, and also of the increasing use of this adaptive strategy with age.

Occurance of "direct hits". Finally, we looked at "direct hits", i.e. estimations that are exactly the same as the actual number of blocks in the grid. It was anticipated that subjects who are applying the ADD strategy will succeed in making such direct hits only on the trials with (very) small quantities, while users of the MIN strategy will produce "direct hits" both on the trials with (very) small and with (very) large numerosities. Therefore, we predicted that the difference in direct hits between the trials with small and large numerosities would become smaller with age.

To test this prediction, we counted for each subject the number of successive direct hits starting from 1 upwards as well as starting from 100 downwards. For instance, when a subject produced direct hits for the trials with one, two, three, four and five blocks but answered with "7" on the trial with six blocks, this subject received a direct hits score of 5 for the domain of small numbers; if the same subject produced direct hits for the trials with 100 and 99 blocks, but missed the one with 98 blocks, he received a direct hits score of 2 for the large numbers domain.

For the small numbers domain, the mean number of consecutive direct hits for the university students, the sixth graders, and the second graders was, respectively, 16.68, 9.25, and 13.3. For the large numbers domain, the mean number of consecutive direct hits was, respectively, 14.05, 9.55, and 3.6. A comparison of the number of direct hits for the trials with small and large quantities for each age group revealed no significant difference for the students and the sixth graders (16.68 vs 14.05 for the university students and 9.25 vs 9.55 for the sixth graders), but for the second graders this difference (13.3 vs 3.6) was significant [$t(9)=-3.1978$; $p<.0109$].

These results provide further evidence for developmental differences in strategy use between the different age groups. However, the finding that the typical second grader produced more than three consecutive direct hits in the domain of large numerosities (i.e., for the trials with 100, 99 and 98 blocks) suggests that some of these young children turned also to the sub-traction procedure at least in the most trivial cases.

Summarizing the results reported so far we can conclude that they yield convincing evidence for the second half of the first hypothesis, namely that the use of the more adaptive

MIN strategy will increase with age; only in the patterns of estimation errors we found no clear evidence for the greater use of the MIN strategy with increasing age. The results are also in line with the first half of the first hypothesis -- stating that the use of the stereotyped ADD strategy will decrease with age --, but here the most striking finding is the unexpectedly low fit of the data with the ADD strategy, even in the youngest age group.

Efficiency of the MIN strategy

To test the second hypothesis, namely that adaptive choices are related to better estimation performance, we calculated correlations between the fits of the subjects' response time and error rate patterns with the segmented curve model, on the one hand, and the following measures of estimation accuracy, on the other:

(a) the subject's mean error rate, i.e. the mean of the absolute deviation of his(her) estimation from the correct answer for the 100 trials, and

(b) the number of responses deviating no more than 10 % from the correct answer.

We remind that the second measure was also used to give feedback about estimation accuracy during the ten example trials which were given before the start of the experiment. Because of this feedback provided during the example trials, we consider the second measure as the best operationalisation of estimation accuracy.

Based on the assumption that making adaptive choices among the addition and subtraction procedure is a crucial aspect of good performance on the estimation task, we expected a negative correlation between the fit of the subjects' response times and error rates with the segmented curve model and the first measure of estimation accuracy, and -- correspondingly -- a positive correlation with the second measure of estimation accuracy.

Response time data. For the response time data, the correlations in the two oldest age groups are in line with the expectations: in the groups of students and sixth graders, the correlations with the first measure of estimation accuracy are -.76 and -.78, respectively, and for the second measure positive correlations of respectively .77 and .81 were found. All four correlations are significant ($p<.05$). For the second graders, on the other hand, no significant correlations were found (.12 for the first measure and .25 for the second measure).

Error rate data. With respect to the error rates, the same pattern emerges. For the adults and the sixth graders, the correlation with the first measure of estimation accuracy is -.50 and -.58, respectively, and for the second measure respectively .47 and .49 were found. Again, all four correlations are significant ($p<.05$). As for the response time data, no significant correlations for the second graders were found (-.11 for the first measure and .52 for the second measure).

In sum, the significant correlations in the two oldest age groups reveal a strong relationship between the extent to which sixth graders and university students applied the MIN strategy, on the one hand, and their estimation accuracy, on the other. For the second graders no such relationship was found. Looking back, one can say that this is not surprising taking into account the rare use of this adaptive strategy in the youngest age group.

DISCUSSION

The present study investigated the development of an adaptive minimization (or MIN) strategy for handling a task in which one has to estimate the numerosity of a number of square blocks in a rectangular 10 x 10 grid. The quintessence of this strategy is that one either repeatedly adds estimations of groups of blocks (= addition procedure) or subtracts the number of empty squares in the rectangular grid from the total amount of squares in that grid (= subtraction procedure), depending on which of both procedures is the most economical one for a given trial.

Fifty subjects (twenty 20-22 year old university students, twenty sixth graders and ten second graders) participated in the experiment. The identification of subjects' estimation strategies was based on the analysis and the mutual confrontation of several kinds of data, namely their responses and response times for all trials in the estimation task and their retrospective reports.

The study aimed at testing two hypotheses. First, we hypothesized that the use of the ADD strategy would decrease and that the use of the MIN strategy would increase with age. Second, the use of the MIN strategy would be associated with better performance on the estimation task.

With respect to the first hypothesis, evidence from different sources supported the expected developmental trend towards the greater use of the MIN strategy. This was not only evidenced by the statistical analysis of the individual response time data using Beem's (1995a, 1995b) program for detecting procedure shifts in cognitive tasks, but also by additional analyses based on subjects' self-reports, systematic estimation errors and direct hits. Only the patterns of error rates revealed no clear evidence for increased use of the adaptive MIN strategy with age. Although the data of the second graders contained the least evidence for the use of the subtraction procedure, we nevertheless found already some undeniable traces of it, restricted, however, to those trials where its superiority was abundantly clear. In terms of Siegler and Shipley's (1995) adaptive strategy choice model, the strategic differences between the distinct age groups documented in the present study are not primarily a matter of acquisition of new procedures for estimating numerosity, but rather of more frequent, more fluent and more adaptive application of one of these procedures, namely the subtraction procedure. As such, this result is in line with the findings of other studies about the role of strategy change in the development of numerical cognition (Lemaire & Siegler, 1995; Newman & Berger, 1984). The present study yielded also some evidence for the first part of the first hypothesis -- namely the decrease of the ADD strategy with age --, but the most important result here was the remarkably low fit of the data with the predictions of the ADD strategy, even in the youngest age group. This means that the hypothetical ADD strategy did not account for the way in which the estimation task was approached in the initial stage of the development of the estimation skill.

With respect to the second hypothesis, we found a strong and positive relationship between the use of the MIN strategy (as evidenced by the individual patterns of response times and error rates), and estimation accuracy in the two oldest age groups. This indicates that making adaptive choices among the different procedures is indeed an important characteristic of good estimators of numerosities. This finding corresponds also to what was found in other studies about the relationship between adaptive strategy choice and learning performance (see, for instance, Geary & Burlingham-Dubree, 1989). The absence of a significant correlation in the group of second graders is probably due to the fact that the MIN strategy was applied only rarely in this age group.

In sum, the results are in line with the findings of other studies using different types of cognitive tasks (Baroody & Ginsburg, 1986; Geary & Burlingham-Dubree, 1989; Hope &

Sherill, 1987; Lefevre, Greenham & Waheed, 1993; Newman & Berger, 1984; Siegler, 1988; Siegler & Jenkins, 1989; Lemaire & Siegler, 1995). These investigations show that (1) subjects make adaptive choices between alternative procedures to minimize cognitive effort in mathematics tasks, (2) the acquisition of a disposition towards adaptive strategy choice is a long and gradual process that begins at a very young age, and (3) there is a strong relationship between the quality of subjects' adaptive choices, on the one hand, and their performance, on the other.

Thus, the present study provides additional empirical evidence for the importance of strategy change to account for developmental and interindividual differences in learning. Moreover, is shows the value of the Beem's (1995a, 1995b) segmented curve model for analyzing adaptive strategy choice in a more detailed way than has been done in previous research, and for relating these strategy choices to performance. However, several theoretical, methodological, and practical questions require further research.

First, the data of this study suggest that a number of subjects have used another estimation procedure than the two procedures that were distinguished in the present rational and empirical analysis, namely the addition and subtraction procedure. Indeed, the graphs of the response time patterns of the university students and the sixth graders suggest that at least some of them applied another estimation procedure, especially for the trials in the middle range. The basic characteristics of this alternative procedure are that it is a very quick one, that it leads to relatively unprecise estimations, and that its duration is not (seriously) affected by the actual numerosity to be estimated (see, for instance, the graphs in Figure 3b and 4b). Probably, it concerns a rather primitive procedure in which the number of blocks is determined instantaneously, as in the well-documented "subitizing" procedure by which small quantities up to $\pm$ 5 are determined (Fuson, 1992). While university students and sixth graders may have applied this primitive, instantaneous estimation procedure for trials in the middle region only, it is plausible that the second graders used it much more frequently, namely for all numerosities that were not countable within the time limit of 20 seconds (see, for instance, the graph in Figure 5b). Further research is needed about the nature of this alternative estimation procedure, its development, its relation with other estimation procedures as well as with estimation

accuracy. This research could profit from additional data-gathering techniques, such as the registration of eye movements (Verschaffel, De Corte, Gielen & Struyf, 1994).

A second important issue is to unravel how users of the MIN strategy choose among the different procedures (see also: Siegler & Jenkins, 1989). According to the rational task analysis, subjects' adaptive choice between the addition and the subtraction procedure for a given trial is based on their own evaluation of the advantage of a particular procedure over another in terms of speed, accuracy, or both. The present study does not provide much information about the processes underlying this evaluation. To unravel these processes we intend to set up new studies in which subjects of different ages will be administered a variety of types of estimation tasks with distinct addition and subtraction procedure eliciting characteristics (e.g., tasks in which the numerosities are presented in square grids of different sizes or even in grids that do not have the form of a square, but a rectangle or a triangle. Under such experimental conditions the efficient utilization of the MIN-strategy puts much higher demands on subjects' mathematical knowledge and self-regulation skills. Besides systematic manipulation of certain task variables, the registration of subjects' eye-movement data while solving the estimation task could yield further information about the mechanisms underlying these adaptive choices. Addition information could also be gained from the administration of separate diagnostic tasks aimed at assessing subjects' mastery of different kinds of conceptual, procedural and metacognitive knowledge that are considered as building blocks for the efficient application of the MIN strategy.

Third, we point to some methodological problems concerning the registration and the analysis of the data, which need a proper solution in future research. A first problem relates to the use of a time limit. In the present study participants had to make their estimation within 20 s. This time limit was inserted as an extra pressure for the subjects to estimate rather than count the numerosities. However, this did not prevent the subjects from counting the smaller numerosities. Moreover, for medium and large numerosities it elicited a considerable number of response times of 20 s., which complicated the regression analyses. In our future work, more effective ways for guaranteeing estimation will have to be built in (for instance, using a reward system that takes into account both the speed and the accuracy of a subject's estimation for a given trial). Another problem deals with the way in which response times were measured. As

said before, this was done by asking subjects to push the space bar, which stopped the clock and made disappear the given numerosity. However, during the individual interviews, some subjects acknowledged that they had pressed the space bar before the actual estimation process was completely finished. This happened especially during estimations involving subtraction: once the number of empty squares was estimated, these subjects pushed already on the space bar (which made the clock stop) and only then subtracted this number from the total number of squares (100 - x = .). This may have resulted in a systematic underestimation of the real response times for the trials that were solved by means of the subtraction procedure. This problem will be solved in future research by asking subjects to respond verbally and by determining their response times based on these verbal responses.

Next there is the statistical technique used for confronting the observed response time and error rate patterns with the hypothetical patterns derived from the ADD and MIN strategy. We used Beem's segmented curve model to analyze the data, since it seemed ideally suited for investigating procedural shifts in cognitive tasks (Beem 1995a, 1995b; Ippel & Beem, 1987). Indeed, this regression seems more appropriate and more informative than other kinds of analyses, such as the categorial approach followed by Newman and Berger (1984) in their investigation of young children's procedures for determining ordinal numbers between 1 and 25. They divided the number range in three distinct groups (i.e. low, middle, and high numbers) for which they calculated the means. The main problem with this approach is that choice of the bin size -- or otherwise stated, the number of subgroups in which the whole number range is divided -- is rather arbitrary, and that the results may differ considerably depending on the bin size chosen. Moreover, Beem's program searches and computes the change point, instead of merely assuming that it is located somewhere in the middle -- an assumption underlying the use of three bins, as in the study of Newman and Berger (1984). Beem's two-phase segmented model is also preferable above a quadratic model, because it allows the detection of two regression equations which do not necessarily need to be continuous in the change point, while this is required in a quadratic model. There are, however, still some problems with the use of the segmented curve model too. The most important problem is that it can only capture one procedure shift, while the visual inspection of the graphs of several participants in the present study suggests the presence of two procedural shifts instead of one.

Finally, we point to some practical implications of the present study. One of the most typical aspects of the ongoing worldwide reform with respect to elementary arithmetic learning and teaching, is the emphasis put on number sense and estimation (National Council of Teachers of Mathematics, 1989; Treffers & de Moor, 1990; Verschaffel & De Corte, 1996). In the current curriculum reform documents for the elementary school more attention is paid at goals like: understanding the (relative) magnitude of numbers, having meaningful benchmarks for discrete and non-discrete quantities, having strategies for counting and estimating (large) quantities, etc. The results of the present study suggest that already at a very young age children are susceptible to learning experiences aimed at the development of number sense and flexible estimation skills. But they also reveal the complex, multifaceted nature of these important aspects of a mathematical disposition. Indeed, the efficient and sophisticated application of the MIN strategy presupposes mastery of a varied and well-elaborated set of mathematical concepts and skills, like understanding of basic arithmetic operations, knowledge of basic addition and multiplication facts, place value concepts, procedures for mental addition and subtraction in the 20-100 range, etc., together with metacognitive knowledge and skills. During the last few years, some efforts have been made to construct instructional materials and techniques for improving children's number sense and estimation skills (Sowder, 1992). In our future work, we intend to investigate the instructional value of a computer-supported learning environment for estimating numerosity based on computerized research tools like the one used in the present study.

REFERENCES

Baroody, A.J., & Ginsburg, H.P. (1986). The relationship between initial meaningful and mechanical knowledge of arithmetic. In J. Hiebert (Ed.), Conceptual and procedural knowledge: The case of mathematics (pp. 75-112). Hillsdale, NJ: Erlbaum.

Beem, A. L. (1995a). A program for fitting two-phase segmented curve models with an unknown change point, with an application to the analysis of strategy shifts in a cognitive task. Behavior Research Methods, Instruments, & Computers, 27, 392-399.

Beem, A. L. (1995b). Segcurve: A program for fitting two-phase segmented curve models with an unknown change point [Program manual]. Leiden, The Netherlands: Leiden University, Centre for the Study of Education and Instruction.

De Corte, E., Greer, B., & Verschaffel, L. (1996). Mathematics teaching and learning. In D. C. Berliner & R. Calfee (Eds.), Handbook of educational psychology (pp. 491-549). New York: Macmillan.

Ericsson, K. A. , & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.

Fuson, K. C. (1992). Research on whole number addition and subtraction. In D.A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 243-275). New York: Macmillan.

Geary, D. C. & Brown, S. C. (1991). Cognitive addition: strategy choice and speed-of-processing difference sin gifted, normal and mathematically disabled children. Developmental Psychology, 27, 398-406.

Geary, D. C., & Burlingham-Dubree, M. (1989). External validation of the strategy choice model for addition. Journal of Experimental Child Psychology, 47, 175-192.

Hope, J.A., & Sherill, J.M. (1987). Characteristics of unskilled and skilled mental calculators. Journal for Research in Mathematics Education, 18, 98-111.

Ippel, M. J., & Beem, A. L. (1987). A theory of antagonistic strategies. In E. De Corte, H. Lodewijks, R. Parmentier, & P. Span (Eds.), Learning and instruction: European research in an international context (pp. 111-121). Oxford: Pergamon.

Lefevre, J.-A., Greenham, S. L., & Waheed, N. (1993). The development of procedural and conceptual knowledge in computational estimation. Cognition and Instruction, 11, 95-132.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. Journal of Experimental Psychology: General, 124, 83-97.

National Council of Teachers of Mathematics (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.

Newman, R. S., & Berger, C. F. (1984). Children's numerical estimation: Flexibility in the use of counting. Journal of Educational Psychology, 76, 55-64.

Schoenfeld, A.H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D.A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 334-370). New York: Macmillan.

Siegler, R. S. (1988). Individual differences in strategy choices. Child Development, 59, 833-851.

Siegler, R. S., & Jenkins, E. A. (1989). How do children discover new strategies? Hillsdale, NJ: Erlbaum.

Siegler, R.S., & Shipley, C. (1995). Variation, selection, and cognitive change. In G. Halford & T. Simon (Eds), Developing cognitive competence: New approaches to process modeling (pp. 31-76). Hillsdale, NJ: Erlbaum.

Sowder, J. (1992). Estimation and number sense. In D.A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 371-389). New York: Macmillan.

Treffers, A., & de Moor, E. (1990). Proeve van een nationaal programma voor het reken-wiskundeonderwijs op de basisschool. Deel 2. Basisvaardigheden en cijferen. [Towards a national mathematics curriculum for the elementary school. Part 2. Basic skills and written computation]. Tilburg: Zwijssen.

Verschaffel, L., & De Corte, E. (1996). Number and arithmetic. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds), International handbook of mathematics education. Part 1 (pp. 99-138). Dordrecht: Kluwer.

Verschaffel, L., De Corte, E., Gielen, I., & Struyf, E. (1994). Clever rearrangement strategies in children's mental arithmetic: a confrontation of eye-movement data and verbal protocols. In: H. Van Luit (Ed.), Research on mathematics learning and instruction in (special) primary schools (pp. 153-180). Doetinchem/Rapallo: Graviant.

Author notes

Lieven Verschaffel is a senior research associate of the Fund for Scientific Research, Flanders (Belgium).

This research was supported by the Research Council of the University of Leuven (G.O.A. grant 93/1). The authors would like to thank Wim Fias for his programming assistance.

Table 1.

Response-time data fits (r^2) of the linear model (Lin) and the segmented curve model (Segm) for the 20 university students. (Also given are the estimated parameters b_1 and b_2 of the segmented curve model and the estimated change point.)

Subj	Lin	Segm	b_1	b_2	Change point
1	0.87	42	234	-42	32
3	0.54	90	297	-411	58
4	0.51	55	297	-94	32
5	0.02	55	166	-158	54
6	1.08	88	379	-520	58
7	0.2	85	343	-379	54
8	1.43	32	59	-98	46
9	0.59	22	10	-339	86
10	0.4	86	356	-551	59
11	5.04	84	215	-313	70
12	1.54	33	28	-286	75
13	0.31	57	292	-129	42
14	1.72	68	131	-355	73
15	0.04	51	168	-172	61
16	0.22	38	164	-140	51
17	0.71	44	55	-331	71
18	0.17	47	93	-391	70
19	0.04	33	60	-306	77
20	2.00E-03	62	173	-156	50
	0.81	56.42			58.89

Table 2.-

Response-time data fits (r^2) of the linear model (Lin) and the segmented curve model (Segm) for the 20 sixth graders. (Also given are the estimated parameters b_1 and b_2 of the segmented curve model and the estimated change point.)

Subj	Lin	Segm	b_1	b_2	Change point
1	0.59	45	52	-425	65
2	1.67	46	35	-22	60
3	5.61	25	24	-393	94
4	2.38	33	74	-580	73
5	2.00E-03	24	501	-13	12
6	3.69	55	109	-448	71
7	0.3	70	187	-344	65
8	2.81	21	19	-494	86
9	14.53	41	6	-271	83
10	6.93	23	18	-494	90
11	1.24	33	7	-489	84
12	0.7	96	354	-430	61
13	0.71	35	16	-449	83
14	3.22	84	366	-493	59
15	1.23	33	304	-31	35
16	0.15	18	297	5	20
17	0.37	32	1133	-24	16
18	0.85	27	25	-473	81
19	3.05	22	27	-458	85
20	2.44	57	76	-81	46
	2.62	41			63.45

Table 3.

Response-time data fits (r^2) of the linear model (Lin) and the segmented curve model (Segm) for the 10 second graders. (Also given are the estimated parameters b_1 and b_2 of the segmented curve model and the estimated change point.)

Subj	Lin	Segm	b_1	b_2	Change point
1	7.43	39	31	86	
2	13.48	24	52	-374	72
3	1.25	31	330	99	49
4	3.15	20	319	35	27
5	1.68	16	-1.116	-761	87
6	10.1	26	47	87	
7	7.44	15	597	32	19
8	2.24	39	329	26	32
9	25.51	46	629	81	32
10	10.87	18	87	98	68
	8.32	27.4			55.9

Figure captions

Figure 1: Visual display of two examples of trials of the estimation task.

Figure 2: Mean response times (in ms) over participants for the three age groups.

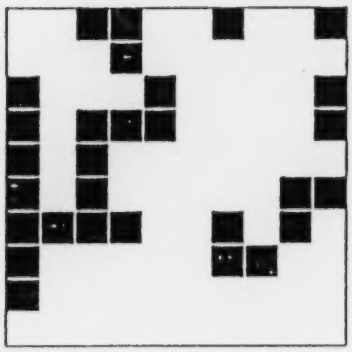
Figure 3: Response times (in ms) results for (a) subject 3 and (b) subject 12 from the group of university students.

Figure 4: Response times (in ms) results for (a) subject 12 and (b) subject 16 from the group of sixth graders.

Figure 5: Response times (in ms) results for (a) subject 4 and (b) subject 1 from the group of second graders.

Figure 6: Mean error rate data over participants for the three age groups.

Figure 1



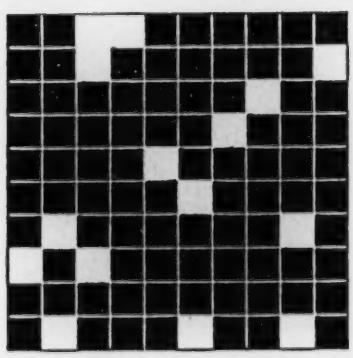
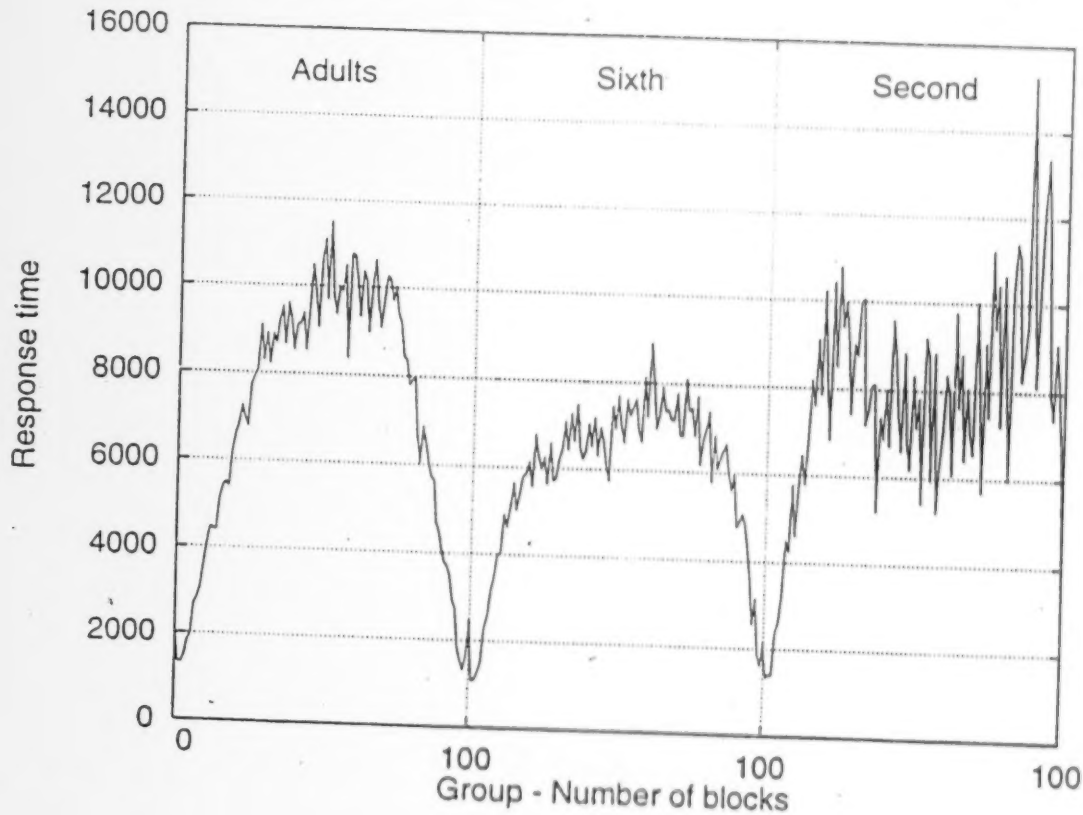
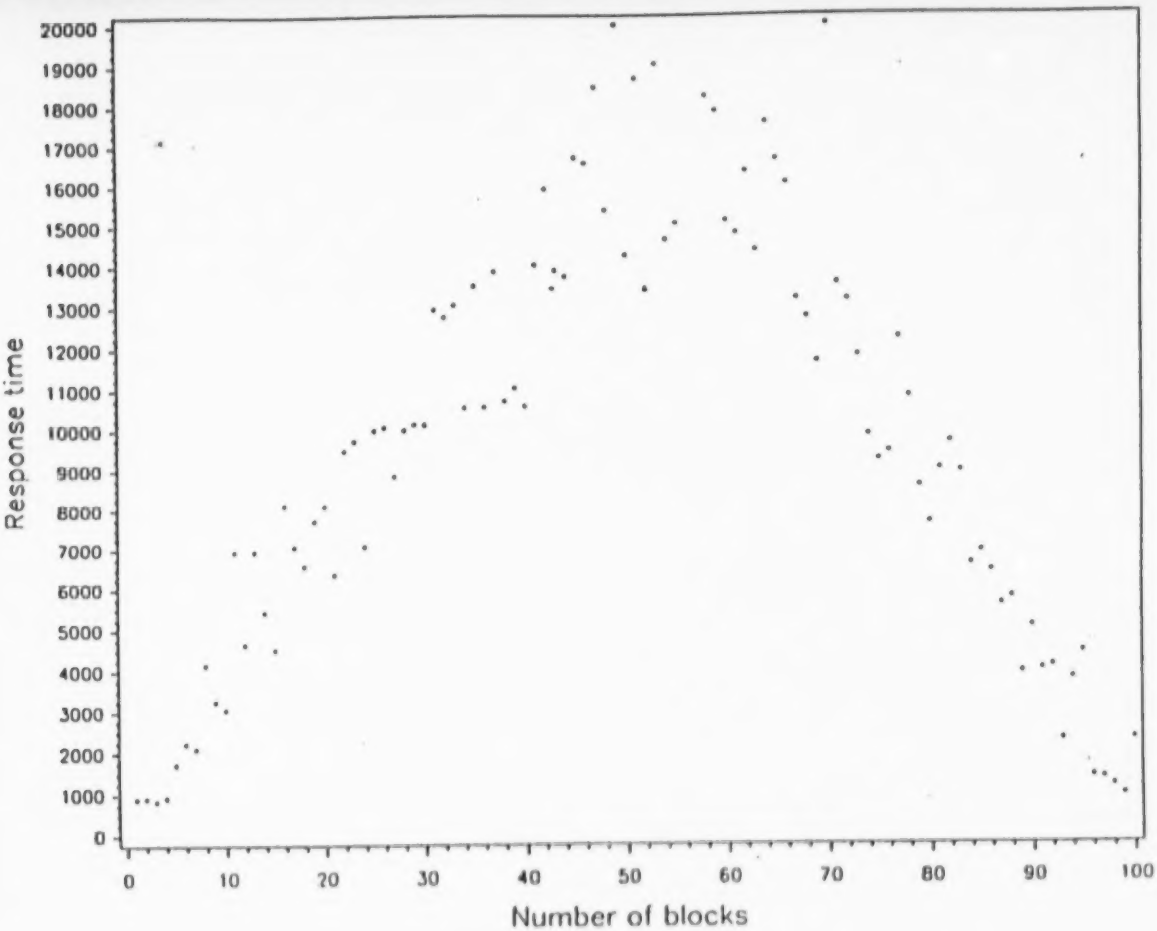
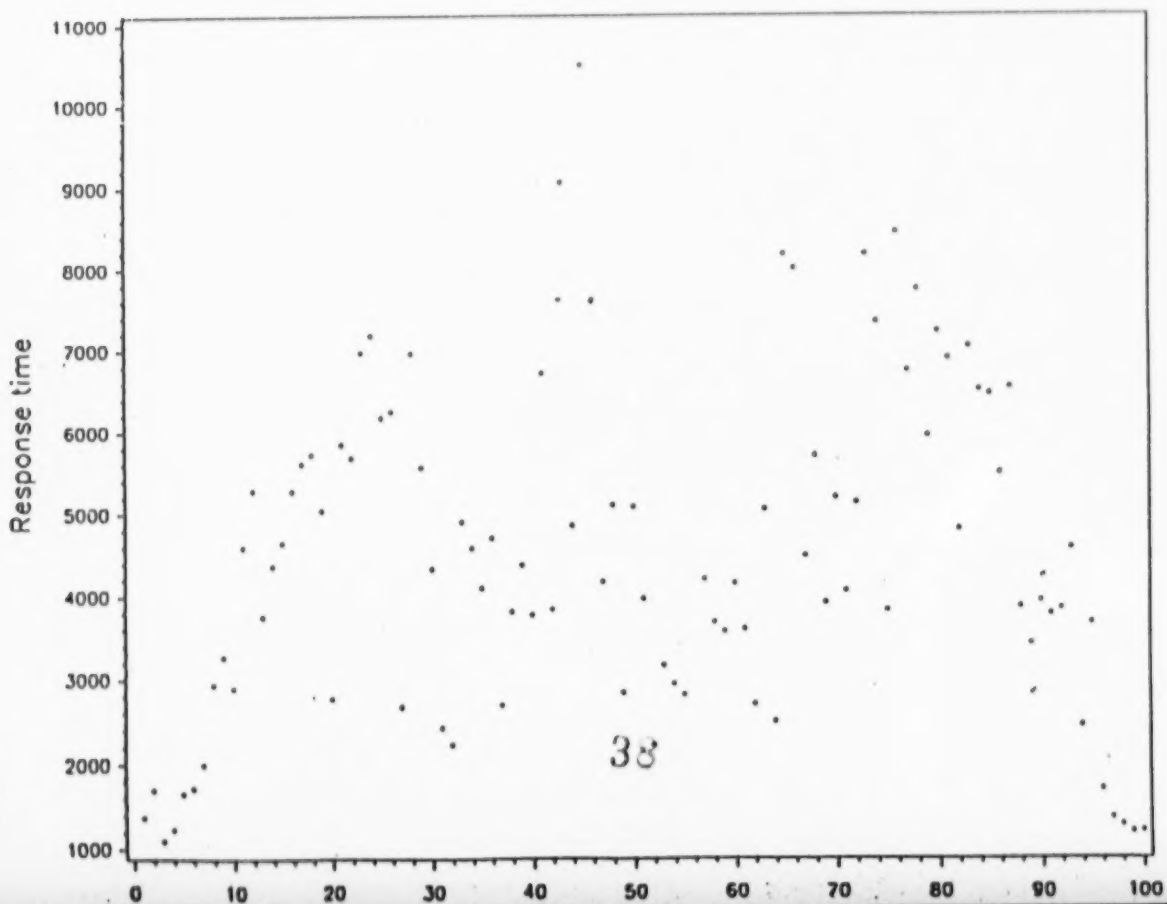
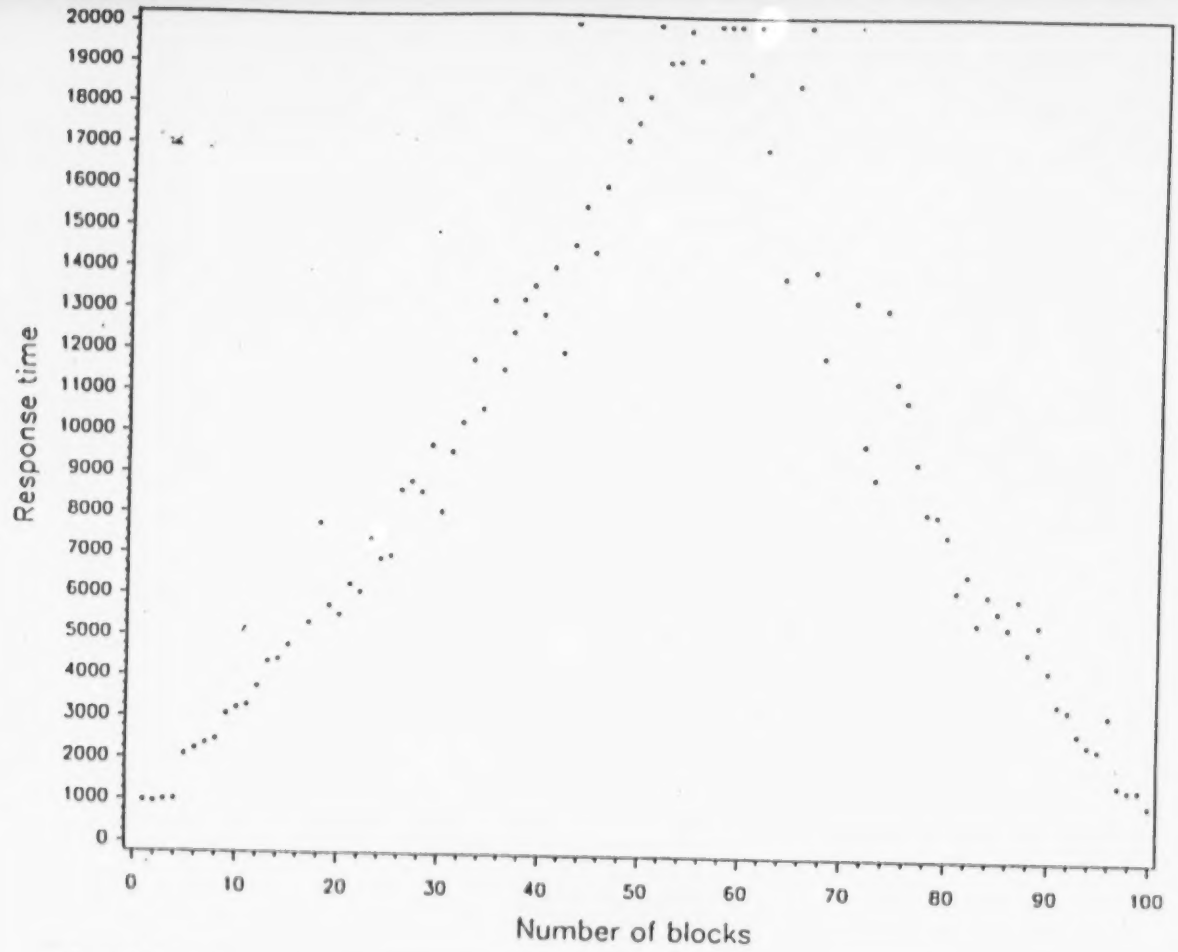


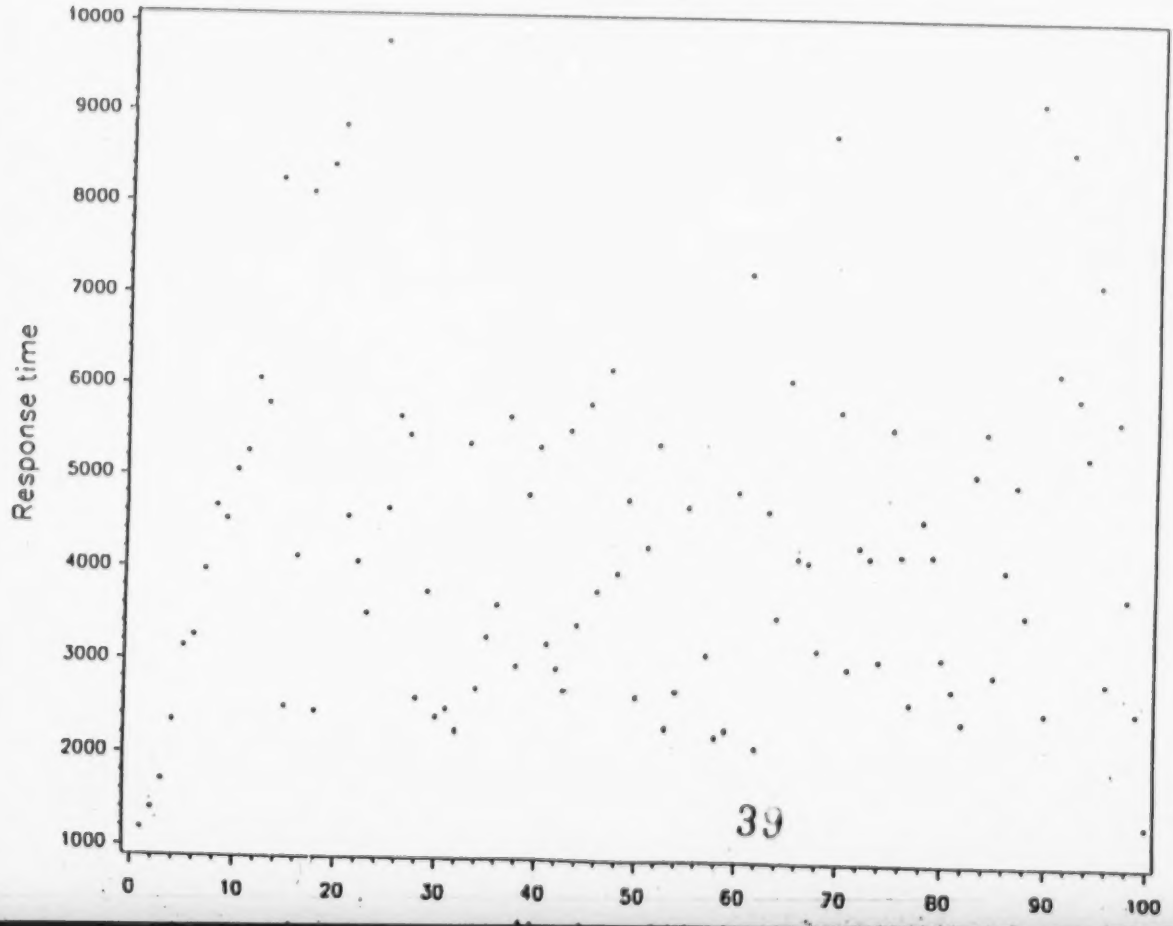
Figure 2

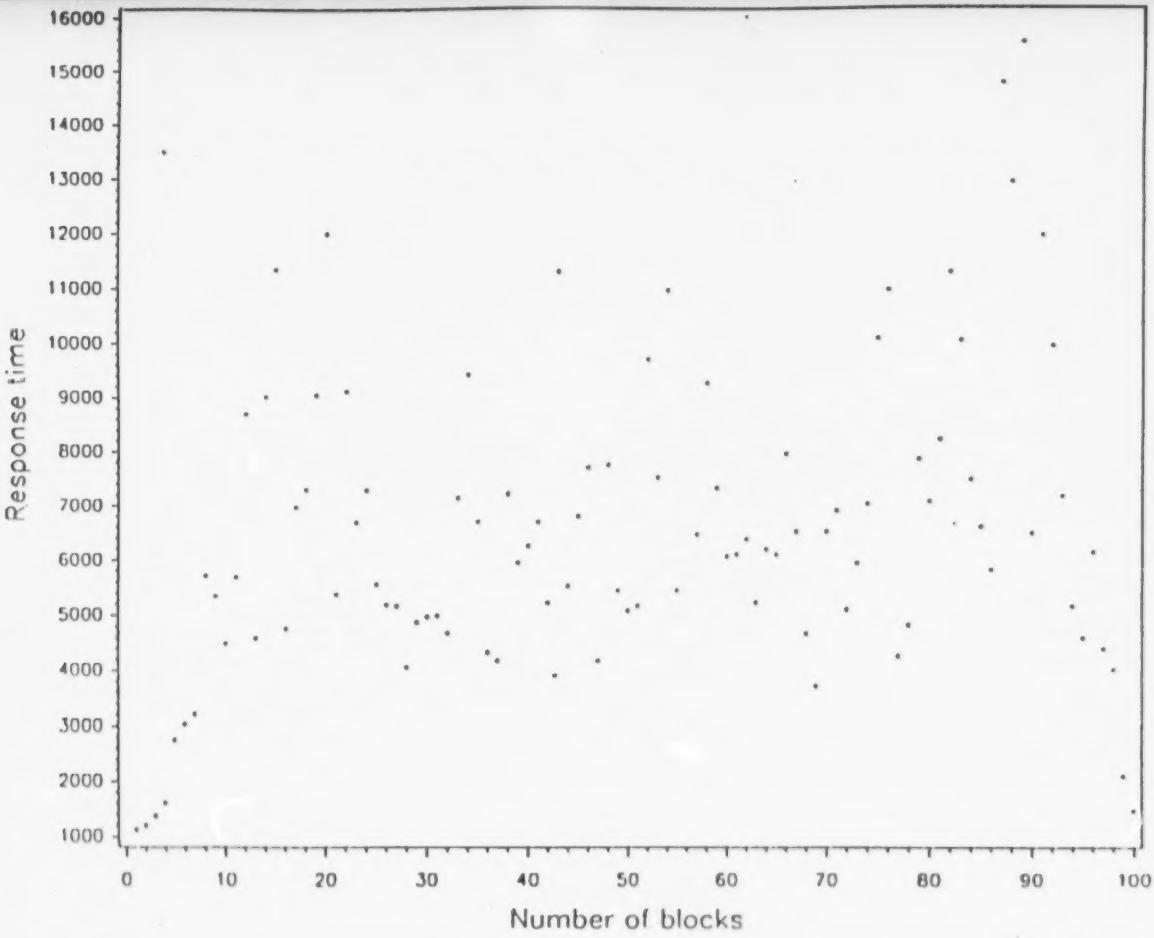

16000
14000
12000
10000
8000
6000
4000
2000
0
Response time
Adults
Sixth
Second
0
100
100
100
Group - Number of blocks


Response time
20000
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
0
10
20
30
40
50
60
70
80
90
100
Number of blocks


Response time
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
10
20
30
40
50
60
70
80
90
100


Response time
20000
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
0
10
20
30
40
50
60
70
80
90
100
Number of blocks


Response time
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
10
20
30
40
50
60
70
80
90
100


Response time
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
10
20
30
40
50
60
70
80
90
100
Number of blocks

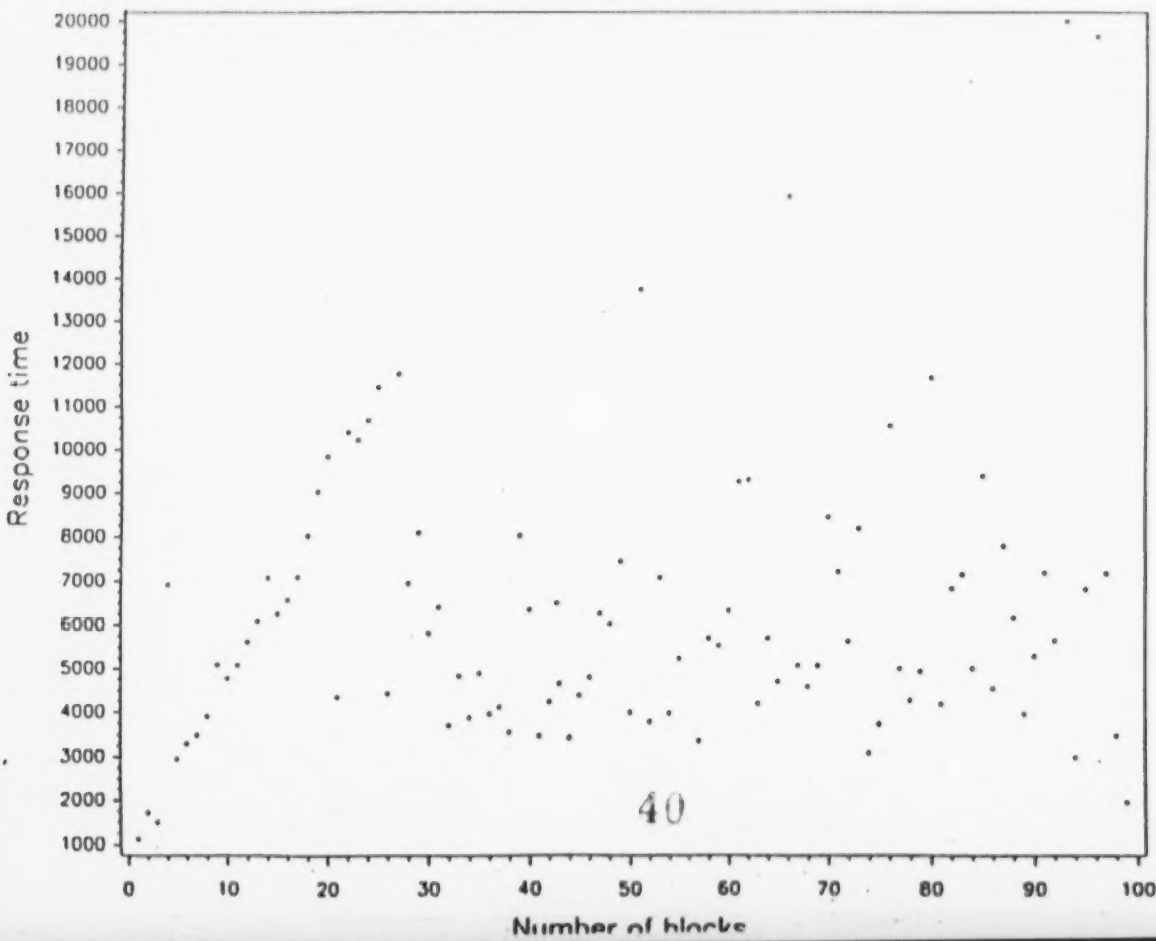
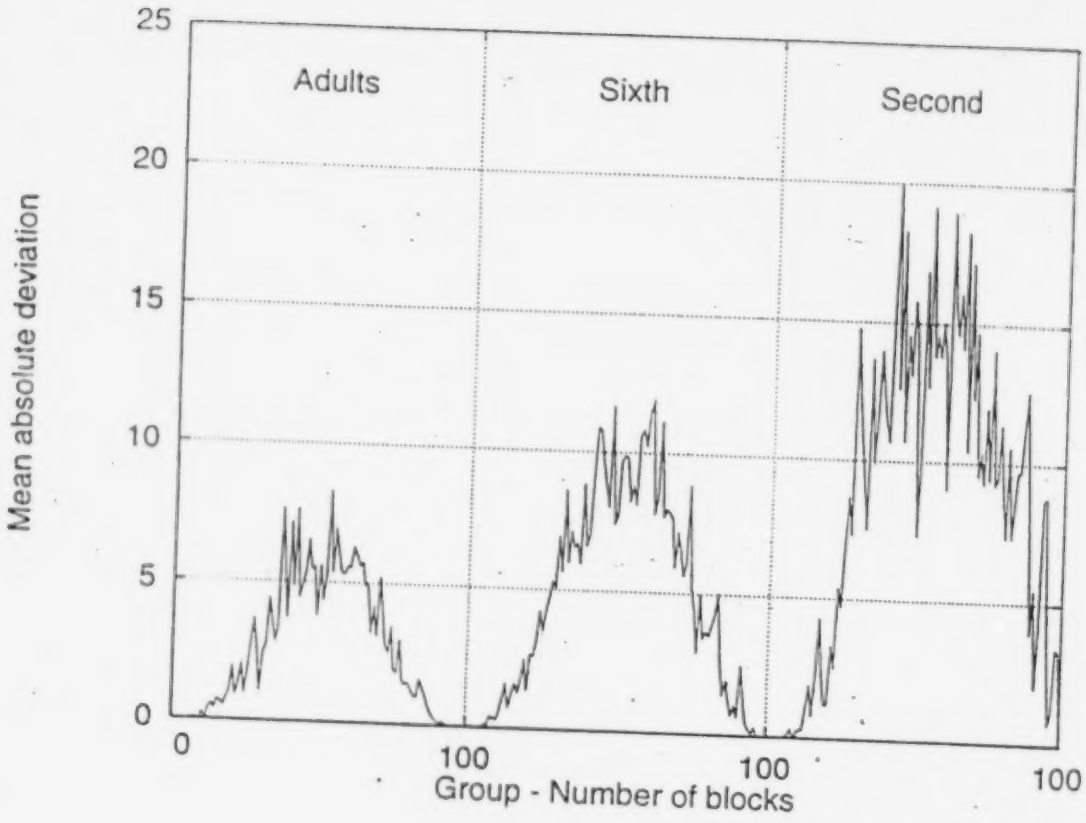

Response time
20000
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
10
20
30
40
50
60
70
80
90
100
Number of blocks

Figure 6


25
20
15
10
5
0
Mean absolute deviation
Adults
Sixth
Second
0
100
100
100
Group - Number of blocks